

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

## **REMARKS**

Claims 1, 33, 36 and 50 are pending in the present application. Applicants have cancelled claims 2-32, 34, 35, and 37-49 and amended claim 33. Support for the amendment of claim 33 is in the original claim 33 and on page 72 of the specification. Thus no new matter has been introduced.

### **Objection to Specification**

The title has been objected to because of the use of the word “novel”. Applicants have amended the title to address the examiner’s concern. As amended the title reads as “Polypeptides and Polynucleotides homologous to Thymosin, Ephrin A Receptors, and Fibromodulin.” Accordingly, the applicants request the objection be withdrawn.

### **Rejections under 35 USC 102(b)**

Claim 1, 33, 36 and 50 are rejected under 35 USC 102(b) as anticipated by Chan et al., the Examiner states that “Chen et al. teaches human eek. Chen et al. provided only a partial sequence. However, the complete sequence of eek is identical to applicant’s SEQ ID NO: 5.” (Office Action at page 2). Applicants traverse.

As rightly pointed out by the Examiner, Chan et al. (discussed as “Chen et al.” in the office action) disclose only a partial, 21 amino acid long sequence that corresponds to residues 706-726 of the full-length polypeptide (GRLAM IVTEYMENGSLDTFLR), which is a translation of exon D2 of human eek nucleotide disclosed in GenBank Accession No. X59291 (See Figure 1 of Chan et al., courtesy copy enclosed herewith). However, the full length polynucleotide, as indicated in the original specification, page 10, was deposited in Genbank September 14, 2000 as accession number NP\_065387.1. The present application is entitled to a priority date of October 18, 1999, based on the filing of Provisional Application U.S.S.N. 60/159,992, which discloses the full length polypeptide of SEQ ID NO: 5 (referred to as AL035703\_A). Filed herewith is Appendix A, a CLUSTALW alignment demonstrates that the amino acid sequence AL035703\_A filed in U.S.S.N. 60/159,992 is 100% identical to the polypeptide of SEQ ID NO: 5 of the instant application. SEQ ID NO: 66, which is the elected polypeptide sequence, is the extracellular region of the full length polypeptide of SEQ ID NO: 5.

Therefore, claim 1 and the dependent claims 33, 36 and 50 are not anticipated by Chan et al. and thus Applicants respectfully request that this rejection be withdrawn.

**Rejections under 35 USC 112, first paragraph**

**Written Description**

Claims 2, 4 and 39 are rejected for failing to comply with the written description requirement as naturally occurring variants are not adequately described. Applicants have cancelled claims 2, 4 and 39 that recited “allelic variants.” Therefore, this rejection is moot and should be withdrawn.

**Enablement**

Claims 27, 28, 39 and 49 are rejected for lack of enablement. The Examiner states that the specification fails to “provide guidance as to the structural, physical, or biological characteristics of polypeptides which differ from the amino acid sequence of SEQ ID NO: 5 by as much as 15%.” (See Office Action, paragraph bridging pages 5 and 6). Claims 27-28, 39 and 49 have been cancelled.

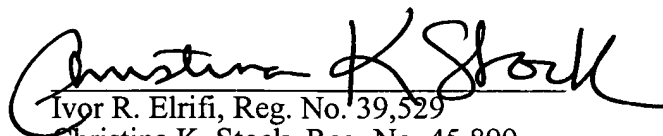
The Examiner also asserts that “the specification fails to provide any guidance as to ‘NOV-associated’ disorders.” (See Office Action, page 7). Without acceding to the propriety of the Examiner’s position, and in order to expedite prosecution, applicants have Applicants have cancelled claims 27, 28, 39 and 49. Therefore, this rejection is moot and should be withdrawn.

### **CONCLUSION**

Applicant submits that the Examiner's rejections have been overcome based on the enclosed amendments and remarks. Applicant therefore respectfully requests that claims 1, 33, 36, and 50 be found allowable at this time. Should any questions or issues arise concerning the application, the Examiner is encouraged to contact Applicant's undersigned attorney at the telephone number indicated below.

Respectfully submitted,

March 15, 2004

A handwritten signature in black ink, appearing to read "Christina K. Stock", is written over a horizontal line.

Ivor R. Elrifi, Reg. No. 39,529  
Christina K. Stock, Reg. No. 45,899  
MINTZ, LEVIN, COHN, FERRIS,  
GLOVSKY and POPEO, P.C.  
Customer Number 30623  
Tel: (617) 542-6000  
Fax: (617) 542-2241

**APPENDIX A: ClustalW alignment of SEQ ID NO: 5 and AL035703 A**

- 1. \_SEQID\_5\_Non\_Prov\_filing**
- 2. AL035703\_A\_Prov\_filing**

**Clustal Details:**

**CLUSTAL W (1.7) Multiple Sequence Alignments**

Sequence type explicitly set to Protein

Sequence format is Pearson

Sequence 1: \_SEQID\_5\_Non\_Prov\_filing 992 aa

Sequence 2: AL035703\_A\_Prov\_filing 992 aa

Start of Pairwise alignments

Aligning...

Sequences (1:2) Aligned. Score: 100

Start of Multiple Alignment

There are 1 groups

Aligning...

Group 1: Sequences: 2 Score:13592

Alignment Score 6364

CLUSTAL-Alignment file created [/opt/curagen/curatool/data/working/ksanth\_24\_clustalwp.align]

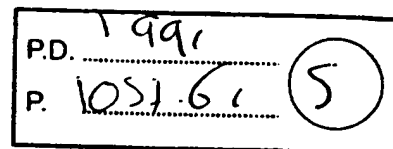
Multiple Alignment:

_SEQID_5_Non_Prov_filing AL035703_A_Prov_filing	MAPARGRLPPALWVVTAAAAAATCVSAARGEVNLDTSTIHGQWGWLTYPAHQWDSINEV MAPARGRLPPALWVVTAAAAAATCVSAARGEVNLDTSTIHGQWGWLTYPAHQWDSINEV
_SEQID_5_Non_Prov_filing AL035703_A_Prov_filing	DESFQPIHTYQVCNWMSPNQNNWLRTSWVFRDQARRVYAEIEFTLRDCNSMPGVLTCKE DESFQPIHTYQVCNWMSPNQNNWLRTSWVFRDQARRVYAEIEFTLRDCNSMPGVLTCKE
_SEQID_5_Non_Prov_filing AL035703_A_Prov_filing	TFNLVYLESDRDLGASTQESQFLKIDTIAADESFTGADLGVRRLKLNTEVRSVGPLSKRG TFNLVYLESDRDLGASTQESQFLKIDTIAADESFTGADLGVRRLKLNTEVRSVGPLSKRG
_SEQID_5_Non_Prov_filing AL035703_A_Prov_filing	FYLAFQDIGACLAILSLRIYYKKCPAMVRNLAAPSEAVTGADSSSLVEVRGQCVRHSEER FYLAFQDIGACLAILSLRIYYKKCPAMVRNLAAPSEAVTGADSSSLVEVRGQCVRHSEER
_SEQID_5_Non_Prov_filing AL035703_A_Prov_filing	DTPKMYCSAEGEWLVPIGKCVCSAGYEERRDACVACELGFYKSAPGDQLCARCPPHSHSA DTPKMYCSAEGEWLVPIGKCVCSAGYEERRDACVACELGFYKSAPGDQLCARCPPHSHSA
_SEQID_5_Non_Prov_filing AL035703_A_Prov_filing	APAAQACHCDLSYYRAALDPPSSACTRPPSAPVNLISSVNGTSTVLEWAPPLDPGGRSDI APAAQACHCDLSYYRAALDPPSSACTRPPSAPVNLISSVNGTSTVLEWAPPLDPGGRSDI
_SEQID_5_Non_Prov_filing AL035703_A_Prov_filing	TYNAYCERCPWALSRCFACGSGTRFVPCQTSLVQASLLVANLLAHMNYSEFWIEAVNGVSD TYNAYCERCPWALSRCFACGSGTRFVPCQTSLVQASLLVANLLAHMNYSEFWIEAVNGVSD
_SEQID_5_Non_Prov_filing AL035703_A_Prov_filing	LSPEPRRAAVVNHITTNQAAPSCQVWVIRQERAGQTSVSLWQEPQPNQIILEYEIKYYEK LSPEPRRAAVVNHITTNQAAPSCQVWVIRQERAGQTSVSLWQEPQPNQIILEYEIKYYEK
_SEQID_5_Non_Prov_filing AL035703_A_Prov_filing	DKEMQSYSTLKAVTTTRATVSGLEKPGTRYVVFQVRARTSAGCGRFSQAMEVETGKPRFRYDT DKEMQSYSTLKAVTTTRATVSGLEKPGTRYVVFQVRARTSAGCGRFSQAMEVETGKPRFRYDT
_SEQID_5_Non_Prov_filing AL035703_A_Prov_filing	RTIVWICLTLLITGLVVLILLICKEKHCQYSKAFQDSDEEKMHYQNGQAPPVFLPLHHP RTIVWICLTLLITGLVVLILLICKEKHCQYSKAFQDSDEEKMHYQNGQAPPVFLPLHHP
_SEQID_5_Non_Prov_filing AL035703_A_Prov_filing	PQKLPEPQFYAEFHTYEEFPGRAGRSTREIEASRIHIEKIIGSGDSGEVGYGRLRVPGQR PQKLPEPQFYAEFHTYEEFPGRAGRSTREIEASRIHIEKIIGSGDSGEVGYGRLRVPGQR
_SEQID_5_Non_Prov_filing AL035703_A_Prov_filing	DVPFAIKALKAGYTEERQRRDFLSSEASIMQGFDPHFNIRLEGWVTEGRLAMIVTEYMNQSD DVPFAIKALKAGYTEERQRRDFLSSEASIMQGFDPHFNIRLEGWVTEGRLAMIVTEYMNQSD
_SEQID_5_Non_Prov_filing AL035703_A_Prov_filing	LDTFLRTHDQQTITMQLVGMRLRGVQAGMRYLSDLGYYVHRDLAARNVLVDSNLVCKVSDFG LDTFLRTHDQQTITMQLVGMRLRGVQAGMRYLSDLGYYVHRDLAARNVLVDSNLVCKVSDFG
_SEQID_5_Non_Prov_filing AL035703_A_Prov_filing	LSRVLEDDPDAAVTTTGGKIPIRWTAPEAIAFRTESSASLVWSFGVVMWEVLAYGERPYW LSRVLEDDPDAAVTTTGGKIPIRWTAPEAIAFRTESSASLVWSFGVVMWEVLAYGERPYW
_SEQID_5_Non_Prov_filing AL035703_A_Prov_filing	NMTNRDVISSVEEGYRLPAFMGC PHALHQLMLDCWHKDRAQRFRFSQIVSVLDALIRSPE NMTNRDVISSVEEGYRLPAFMGC PHALHQLMLDCWHKDRAQRFRFSQIVSVLDALIRSPE
_SEQID_5_Non_Prov_filing AL035703_A_Prov_filing	SLRATATVSRCPPPAFVRSCFDLRGGSGGGGGLTVGDWLDSIRMGRYRDHFAAGGYSSLG SLRATATVSRCPPPAFVRSCFDLRGGSGGGGGLTVGDWLDSIRMGRYRDHFAAGGYSSLG
_SEQID_5_Non_Prov_filing AL035703_A_Prov_filing	MVLRMNAQDVRAIGITLMGHQKEILGSIQTME MVLRMNAQDVRAIGITLMGHQKEILGSIQTME

# *eek* and *erk*, new members of the *eph* subclass of receptor protein-tyrosine kinases

Joanne Chan & Valerie M. Watt

Department of Physiology, University of Toronto, Toronto, Ontario, Canada M5S 1A8



We have identified human and rat DNAs encoding two novel members of the *eph* subclass of putative receptor protein-tyrosine kinases. Rat cDNA clones encoding *eek* (*eph*- and *elk*-related kinase) were isolated from a brain cDNA library probed with DNA encoding the kinase region of the insulin receptor-related receptor. The predicted *eek* protein contains all the amino acid residues conserved in the catalytic domains of protein-tyrosine kinases and is most similar to two putative receptor protein-tyrosine kinases of the *eph* subclass, *elk* (69%) and *eph* (57%). Human genomic DNAs encoding part of *eek* (*EEK*) as well as another putative protein-tyrosine kinase most similar to *elk* (90%), *ERK* (*elk*-related kinase), were isolated and partially characterized. The novel identity of these two *eph*-family genes was further supported by Southern blot analyses and localization to human chromosome 1. In Northern blot analysis of rat RNA, DNAs encoding rat *eek* and human *ERK* hybridized to transcripts most abundant in brain and lung, respectively. These two new members of the *eph* subclass of receptor protein-tyrosine kinases, *eek* and *erk*, may therefore have tissue-specific functions distinct from those of other *eph* family members.

## Introduction

Protein-tyrosine kinases (PTKs) are structurally and functionally related enzymes intimately involved in signal transduction. Initially discovered as transforming proteins of acutely oncogenic retroviruses (Hunter & Cooper, 1985), altered versions of cellular PTKs have since been implicated in the etiology of certain human malignancies (e.g. Konopka *et al.*, 1984; Martin-Zanca *et al.*, 1986). Under physiological conditions, some PTKs function as receptors for a variety of hormones and growth factors to alter such diverse cellular processes as metabolism, growth and differentiation (Yarden & Ullrich, 1988). Ligand binding to the extracellular region of receptor PTKs somehow activates the cytoplasmic catalytic domain to phosphorylate specific substrates such as the GTPase activating protein (GAP, Kazanietz *et al.*, 1990), phospholipase C (Meisenhelder *et al.*, 1989) and phosphatidylinositol 3-kinase (Auger *et al.*, 1989).

Receptor PTK subclasses, defined on the basis of structural similarity (Hanks *et al.*, 1988; Yarden & Ullrich, 1988; Ullrich & Schlessinger, 1990), include those of the epidermal growth factor receptor (EGFR), the insulin receptor (IR), and the platelet-derived and

fibroblast growth factor receptors (PDGFR, FGFR). Within these subclasses are putative receptor PTKs whose presumptive ligands are unknown (Hanks *et al.*, 1988; Ullrich & Schlessinger, 1990). Insight into the function of these putative receptors will be facilitated if, as expected, family members that exhibit limited divergence play similar roles in cellular physiology (Hanks *et al.*, 1988). It is likely, however, that the elucidation of the functional roles of these PTKs will continue to be outpaced by their rate of discovery.

The *eph* PTK defined a new receptor PTK subclass (Hirai *et al.*, 1987) which on the basis of structural similarity also includes *elk* (Letwin *et al.*, 1988). The *eph* full-length cDNA predicts a transmembrane receptor PTK featuring a single Cys-rich region in the extracellular domain and an uninterrupted PTK domain (Hirai *et al.*, 1987). Two lines of evidence suggest that *eph* may be involved in oncogenesis: *eph* is overexpressed in several human carcinomas (Hirai *et al.*, 1987; Maru *et al.*, 1988); and overexpression of the *eph* gene enabled NIH3T3 cells to form tumors in nude mice and colonies in soft agar (Maru *et al.*, 1990).

We report the isolation and characterization of rat cDNA clones encoding a novel PTK, *eek*, whose predicted amino acid sequence within its kinase domain exhibits extensive similarity with the sequence of receptors belonging to the *eph* subclass of PTKs. Human DNAs encoding *eek* (*EEK*) as well as another novel member of the *eph* subclass, *erk*, have been isolated and used to localize both the *EEK* and *ERK* genes to human chromosome 1. The tissue-specific expression of these two new members of the *eph* subclass of receptor PTKs is distinct from other known members of the *eph* family; *eek* expression is brain-specific and *erk* mRNA is most abundant in lung.

## Results and discussion

To identify previously unknown PTKs, we used a DNA probe encoding the kinase region of the insulin receptor-related receptor (IRR, Shier & Watt, 1989) to screen at reduced stringency a rat brain cDNA library. PTKs are abundant in the central nervous system and have been implicated in such brain-specific functions as myelination and neuronal differentiation (Nairn *et al.*, 1985; Edwards *et al.*, 1988). Nucleotide sequence analysis of the entire insert DNA (867 bp) of one cDNA clone which hybridized with the IRR probe, *λ*reek.18, revealed a single open reading frame encoding 289 amino acids. Subsequent screening of another rat brain cDNA library with the insert DNA of *λ*reek.18 identified an overlapping clone, *λ*reek.32, that extended the sequence 3' by 249 nucleotides to a stop codon and a further 1.7 kb to a putative polyadenylation signal. The

Correspondence: V.M. Watt, Department of Physiology, University of Toronto, Medical Sciences Building, Room 3360, 1 King's College Circle, Toronto, Ontario, Canada M5S 1A8

Received 7 May 1990; accepted in revised form 4 February 1991

		▼D1	
reek	RIHIEKIIGSGESGEVVCYGRQLQVPGQRDVPV	<b>AIKALKAGYTERQR</b>	45
relk	FVK..EV..A..F...YK...KL..K.EIY....T.....S.K..		58
heph	WLMVDTV..E..F...YR.T.RL.S.DCKT....T..DTSPGG.W		675
herk		..T..S....KN.	13
		▼D2	
heek		.....T	17
reek	QDFLREAAIMGQFDHPNIIIRLEGVVTRGRLAMIVTEYMENGLDA		90
relk	R...S..S.....KS.PV..I..F....A..S		103
heph	WN.....T.....S..H.LH.....KRKPI..I..F...AA...		720
herk	R...S..S.....V.H.....KSTPV..I..F.....S		58
		▼D3	
heek	...		20
reek	FLRTHDGQFTILQLVGMLKGVGAGMYLSDLGYIHRDLAARNILV		135
relk	...QN.....VI.....R.IA...K...EMN.V.....		148
heph	...ERED.LVPG...A..Q.IAS..N...NHN.V.....		765
herk	...		61
		▼D4	
reek	DGRLVCKVSDFGLSRALEDD-PEAAYTTA-GGKIPIRWTAPEAIA		178
relk	NSN.....Y.Q..TSDPT..SSL.....V.....		193
heph	NQN.C.....T.L.D.--FDGT.E.Q-.....		807
		▼D5	
reek	FRTFSSASDVWSFGVVMWEVLAYGERPYWNMTNQDVISSVEEGYR		223
relk	Y.K.T.....Y.I.....MSF.....D.S.....NAI.QD..		238
heph	H.I.TT.....I.....SF.DK..GE.S..E.MK.I.D...		852
reek	LPAPMGCPRALHQIMLDCWHKDRARPRFSHVVSLEALVHSPES		268
relk	..P..D..A.....Q...NS....AEI.NT.DKMIRN.A.		283
heph	..P.VD..AP.YE..KN..AY...R..H.QKLOAH..Q.LAN.H.		897
reek	LRATATVSRCPA-PAFARSCFDLR--AGGNGNGDLTVGDWLD SIR		310
relk	.KTV..ITAV.SQ.LLD...--IPDFT.FT-----D...SA.K		320
heph	..TI.NFD-----P.VTLR.PSL.SD.IPYR..SE..E...		915
reek	MGRYRDHFAAGGYSSGLMVLHMNAQDVRA LGITLMGHQKKILGSI		355
relk	.VQ...S.LTA.FT..QL.TQ.TSE.LLRI.V..A.....S..		366
heph	.K..IL..HSA.LDTMEC..ELT.E.LTQM....P....R..C..		981
reek	QTMRSQLSCTQGPRRI.		372
relk	HS..VQMNQSPSVMA		380
heph	.GFKD		984

Figure 1 Predicted amino acid sequences of rat *reek*, human *reek*, and human *erk*, aligned with those of the closely related rat *elk* (Letwin *et al.*, 1988) and human *eph* (Hirai *et al.*, 1987). The rat *reek* sequence numbered from 1 is from *reek*.18 with the addition of the 3'-most 83 amino acids from *reek*.32. Sequences were aligned using the GAP programme from the University of Wisconsin Genetics Computer Group Sequence Analysis Software Package. Bold letters represent amino acids conserved among kinases (Hanks *et al.*, 1988). Dots replace residues which are identical to the corresponding residues in the rat *reek* sequence. Hyphens represent gaps introduced in sequences to maximize alignment. Triangles demarcate exons D1 to D5 of the *eph* gene (Maru *et al.*, 1988): closed, conserved among *eph* family members where known; open, not conserved. The nucleotide sequences for rat *reek*, human *EEK* and human *ERK* have been submitted to the GenBank™/EMBL Data Bank with accession numbers X59290, X59291, and X59292.

composite predicted protein (Figure 1) contains all the amino acid residues conserved in the catalytic domains of PTKs (Hanks *et al.*, 1988) including the potential ATP binding site (Gly<sup>9</sup>-X-Gly-X-X-Gly<sup>14</sup> and Lys<sup>34</sup>). In addition, two sequences (Asp<sup>127</sup>-Leu-(Ala-Ala-Arg)-Asn<sup>132</sup> and Pro<sup>168</sup>-Ile-Arg-Trp-Thr-Ala-Pro-Glu<sup>175</sup>, Figure 1) specifically conserved in tyrosine rather than serine/threonine kinases as well as a potential phosphorylation site, Tyr<sup>160</sup>, at a position analogous to the major autophosphorylation site in pp60<sup>src</sup> (Smart *et*

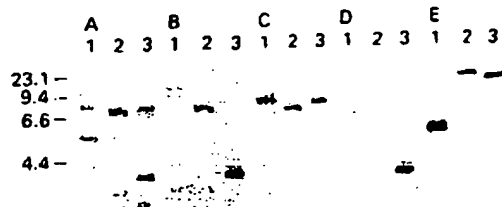
*al.*, 1981), are also present. A computer search of sequence databases (EMBL, GenBank and SWISS-PROT, December 1990) revealed that we had identified a novel protein that exhibits striking amino acid similarity in its kinase domain to members of the *eph* subclass of receptor PTKs, *elk* and *eph* (69% and 57% identity, respectively; also see Figure 1). This putative PTK is less similar to PTKs of other receptor as well as non-receptor subclasses: ~32% to 34% identity to representative members of the IR, EGFR, PDGFR, and



FGFR subclasses; and ~40% to 43% with those of the *src*, *abl*, and *fps/fes* subclasses (Figure 1; Hanks *et al.*, 1988; Kornbluth *et al.*, 1988). The similarity between *EEK* and the other *eph* family members, *eph* and *elk*, also extends into the carboxy-terminal tail (~43% identity, also see Figure 1), the region thought to exert negative control over receptor PTK signalling function (Ullrich & Schlessinger, 1990). Therefore, we have named this novel putative PTK *EEK*, for *eph*- and *elk*-related kinase.

Southern blot analysis of human genomic DNA revealed that a rat *EEK* cDNA probe hybridized at reduced stringency to multiple fragments in each digest (Figure 2A), suggesting that this rat *EEK* probe could identify several *EEK*-related human DNA sequences. At the highest stringency at which any hybridization was observed with the rat *EEK* cDNA probe, two human fragments were detected in each digest (Figure 2B). Hybridization with a rat *elk* cDNA probe (Letwin *et al.*, 1988) indicated that one of these fragments encoded the human homologue of rat *elk* (Figure 2C). To confirm that the other hybridizing fragment was the human homologue of rat *EEK*, we used rat *EEK* cDNA as probe to isolate part of the human *EEK* gene (*EEK*) from a human genomic library. The region of human *EEK* homologous to the rat *EEK* cDNA probe hybridized selectively to the fragments detected under high stringency by rat *EEK* DNA (Figure 2B, D) that had not hybridized with rat *elk* DNA (Figure 2C).

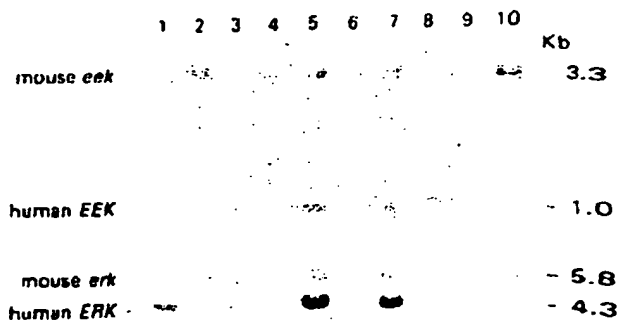
Nucleotide sequence analysis of an ~1 kb fragment of human *EEK* genomic DNA that hybridized with the rat *EEK* cDNA probe identified an exon which exhibits high identity with the rat *EEK* cDNA (95% amino acid, 92% nucleic acid, Figure 1) and much less similarity with its closest known relative, *elk* (70% amino acid, 71% nucleic acid, Letwin *et al.*, 1988). This *EEK* exon corresponds to amino acid residues 74 to 93 of the rat *EEK* cDNA (Figure 1) and to the analogous kinase domain exon D2 of the *eph* gene (Maru *et al.*, 1988). The position of both intron/exon junctions of *EEK* exon D2 (gcccgccagGC...GGgtgcgt), which are similar to the consensus acceptor and donor splice sites (Breathnach & Chambon, 1981), are conserved between *EEK* and the *eph* gene. Other genes encoding PTKs of the same subclass, such as the *src* (Maru *et al.*, 1988) and the IR (Shier & Watt, 1989) subclasses, have also been reported to exhibit conserved exon/intron organization throughout the entire kinase domain.



**Figure 2** Southern blot analysis of genes encoding members of the *eph* family. Human genomic DNA was digested with *Eco*RI (lane 1), *Hind*III (lane 2) or *Bgl*II (lane 3). Filters were hybridized with a rat *EEK* cDNA probe (rEEK.D12p, a 246 bp *Pst*I fragment encoding amino acid residues 22 to 101, Figure 1) and washed at 42°C (A) or at 60°C (B); with a rat *elk* cDNA probe (nucleotides 1-308, Letwin *et al.*, 1988; encoding amino acid residues 1-102, see Figure 1) (C); with human *EEK* genomic DNA (hEEK.D2p, an ~1 kb *Pst*I fragment encoding D2) (D), or with human *ERK* genomic DNA (hERK.D12c, an ~400 bp *Eco*RI fragment encoding part of D1 and all of D2) (E) and washed at 60-62°C.

We also isolated another recombinant phage that contained human DNA which hybridized selectively to human genomic fragments (Figure 2E) detected only at reduced stringency by the rat *EEK* cDNA probe (Figure 2A). Analysis of the nucleotide sequence of this human isolate revealed that it was most closely related to the *eph* subclass member, *elk* (Letwin *et al.*, 1988); we have named it *erk* for *elk*-related kinase. Over the coding region sequenced, the human *ERK* fragment exhibited high identity (90% amino acid, 81% nucleic acid) with the rat *elk* cDNA (Letwin *et al.*, 1988) and lower identity with its next closest known relative, rat *EEK* (74% amino acid, 76% nucleic acid, Figure 1). This human *ERK* genomic fragment (~400 bp from the linker to an internal *Eco*RI site) contains a single exon corresponding to part of exon D1 as well as all of exon D2 of the *eph* gene (Maru *et al.*, 1988). The predicted splice junction at the end of exon D2 (CGGgtaggg) is similar to the consensus donor splice site (Breathnach & Chambon, 1981). The lack of an intron between exons D1 and D2 in the *ERK* gene was somewhat unexpected given that this intron is conserved between the genes encoding both *erk* (Figure 1) and *eph* (Maru *et al.*, 1988). Possibly, an intron was lost in *ERK* as a result of reverse transcription of a partially processed pre-mRNA that was re-inserted downstream from a promoter sequence. A similar mechanism of intron loss has been implicated in the rat *proinsulin* I gene (Soares *et al.*, 1985).

We have used genomic DNAs from the human *EEK* and *ERK* genes and from 14 human-mouse somatic cell hybrids to localize *EEK* and *ERK* within the human genome. A DNA probe which hybridized specifically with the *EEK* gene, hEEK.D2p, detected a single *Pst*I fragment only in hybrids containing human chromosome 1 (1.0 kb, Figure 3). Similarly, the hERK.D12c probe which hybridized specifically with the *ERK* gene, detected a single *Pst*I human fragment in the same hybrids (4.3 kb, Figure 3). As expected, mouse-specific fragments which hybridized to hEEK.D2p and hERK.D12c were present in all hybrids (3.3 kb and 5.8 kb, respectively, Figure 3). Among all 14 hybrids, chromosome 1 showed 100% concordance with *EEK*



**Figure 3** Southern blot analysis of the human *EEK* and *ERK* genes in somatic cell hybrids. Genomic DNA from human (lane 1), mouse (lane 2) and human-mouse hybrid (lanes 3-10) cell lines was digested with *Pst*I and hybridized with human *EEK* DNA (hEEK.D2p, top) or with human *ERK* DNA (hERK.D12c, bottom). Human chromosome 1 and the human DNA fragments hybridizing to *EEK* (1.0 kb) or to *ERK* (4.3 kb) are concordantly present (lanes 3, 5, 7) or absent (lanes 4, 6, 8-10).

Table 1 Segregation of human *EEK* and *ERK* sequences with human chromosomes in somatic cell hybrids

Gene	Chromosome*	Human chromosome†																							
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	X	Y
+	+	6	2	2	3	2	2	4	3	0	4	2	3	2	4	3	1	1	4	3	4	4	0	5	1
+	-	0	3	4	3	4	4	2	3	6	2	4	3	4	2	3	5	5	2	3	2	2	6	1	5
-	+	0	3	2	4	4	2	5	3	0	4	1	4	1	5	2	1	3	2	3	4	4	4	5	0
-	-	8	5	6	4	4	6	3	5	8	4	6	4	7	3	6	7	5	6	5	3	4	3	0	8
% discordancy (n = 14)		0	46	43	50	57	43	50	43	43	43	38	50	36	50	36	43	57	29	43	46	43	77	55	36

\* Chromosome scored '+' if present in greater than 10% of mutants.

\* Chromosome scored '+' if present in greater than 10% of metaphases scored

† Chromosomes were not scored if translocations were present in human parental cells

and *ERK* whereas all other chromosomes were excluded by at least 29% discordancy (Table 1).

Our localization of both the *EEK* and *ERK* genes to chromosome 1 demonstrates that these genes map to a chromosomal location distinct from that of the closely related *eph* gene present on human chromosome 7 (Maru *et al.*, 1988). In addition, it raises the possibility that *EEK* and *ERK* may have arisen by duplication of an ancestral gene. A similar gene duplication event has been suggested previously to have given rise to the genes encoding the  $\beta$  type PDGFR and *c-fms*: both have been shown to be on the same chromosome in the human and mouse genomes, and to be tandemly linked in the human genome (Buchberg *et al.*, 1989; Roberts *et al.*, 1988).

The tissue distribution of mRNA which hybridized with DNA encoding the two novel *eph* subclass members was assessed by Northern blot analysis (Figure 4). To ensure detection of only *eek* or *erk* transcripts, we used DNA probes and hybridization conditions which detected only single fragments on Southern blots of rat genomic DNA (data not shown). A rat *eek*

cDNA probe containing part of the 3'-untranslated region detected hybridizing transcripts only in rat brain (Figure 4). Even with prolonged exposure, hybridization was not detected in other tissues although multiple, larger transcripts were weakly detected in brain (data not shown). In contrast, a human *ERK* genomic probe hybridized to transcripts that were most abundant in lung (Figure 4) and that were also detected on prolonged exposure in placenta, brain, and kidney (data not shown). Hybridization with rat  $\beta$ -actin DNA (Nudel *et al.*, 1983) verified that each lane had approximately the same amount of mRNA and that the mRNA was intact (Figure 4). These tissue distributions of *eek* and *erk* RNAs differ from those of other known members of the *eph* subclass: *elk* mRNA is present in testis as well as in brain (Letwin *et al.*, 1988); *eph* mRNA in kidney, testis, liver as well as in lung (Maru *et al.*, 1988). Also, the sizes of the *eek* and *erk* transcripts (both larger than 28S rRNA, see Figure 4) were larger than those for *eph* and *elk*, large enough to potentially encode an extracellular binding domain as well as the PTK catalytic domain.

A comparison of the predicted amino acid sequences of both *eek* and *erk* (Figure 1) suggests that *eek* and *erk* may be new receptor-type PTKs of the *eph* subclass. Primary structures for several members of the *eph* subclass have been predicted to contain potential ligand binding regions (Hirai *et al.*, 1987; Lindberg & Hunter, 1990; Lhotak *et al.*, 1991), although the ligands which bind these putative receptors are currently unknown. Since oncogenic involvement has been implicated for the putative *eph* receptor PTK (Hirai *et al.*, 1987; Maru *et al.*, 1988, 1990), *eek* and *erk* also may play roles in certain types of neoplastic transformation.

## Materials and methods

### Library screening, cloning and sequencing

A random primed rat brain cDNA library constructed in  $\lambda$ gt11 (Auld *et al.*, 1988) was screened with an ~1 kb BamHI fragment encoding part of the kinase domain of guinea pig IRR (residues 1058 to 1194, Shier & Watt, 1989). Insert DNA from one isolate,  $\lambda$ eek.18, was then used to screen another rat brain cDNA library (Clontech) as well as a human genomic library in  $\lambda$ Charon 4A (Lawn *et al.*, 1978). In each library screen, duplicate nitrocellulose filters were hybridized in 30% formamide at 42°C (Wahl *et al.*, 1979) with DNA probes labelled with [ $\alpha^{32}$ P]dCTP (Feinberg & Vogelstein, 1983) and washed in 15 mM sodium chloride, 1.5 mM sodium citrate and 0.1% sodium dodecyl sulfate at 42°C. Insert DNAs from  $\lambda$ eek.18,  $\lambda$ eek.32,  $\lambda$ eek.23 and  $\lambda$ erk.7 were subcloned into vectors pEMBL18 or 19 (Allison *et al.*, 1985) and deletions were created by restriction endonuclease digestion. Single-stranded DNA templates were sequenced by the dideoxy

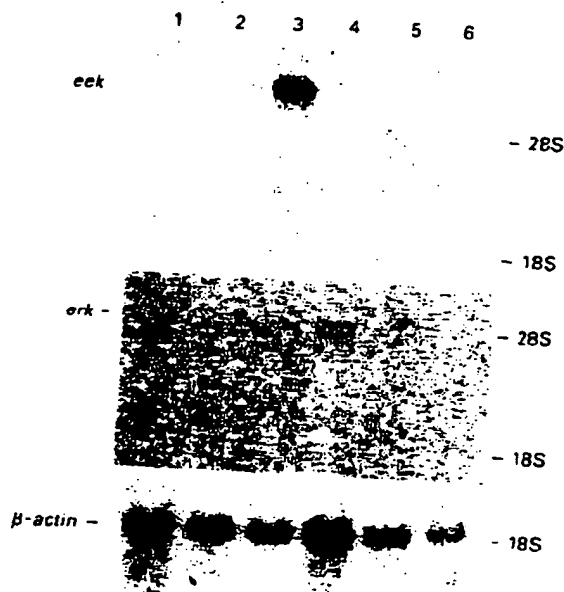


Figure 4 Northern blot analysis of the tissue distribution of *eek* and *erk* transcripts. Poly(A) RNA (2  $\mu$ g) was extracted from rat small intestine (lane 1), placenta (lane 2), brain (lane 3), lung (lane 4), kidney (lane 5), and testis (lane 6). After gel electrophoresis and transfer, the mRNA was hybridized with a rat *eek* cDNA probe ( $\lambda$ eek.XE, an ~1.5 kb fragment between ~200 bp 3' to the stop codon and the linker; top), human *ERK* genomic DNA ( $\lambda$ ERK.D12e, middle), or  $\beta$ -actin DNA (Nudel *et al.*, 1983; bottom); and then washed at 55°C.

method using the Klenow fragment of DNA polymerase (Sanger *et al.*, 1977), with ambiguities resolved using the modified T7 DNA polymerase and dITP (Tabor & Richardson, 1987; Sequenase, USB).

#### Southern and Northern blot analysis

Genomic DNAs from human leukocytes, cultured human fibroblasts, mouse fibroblasts or human-mouse somatic cell hybrids (Shier *et al.*, 1990; Watt & Willard, 1990) were digested with restriction endonucleases and size fractionated on 1% agarose gels before transfer to filters (Southern, 1975; Towbin *et al.*, 1979). Poly (A) RNA (2 µg), extracted using guanidine thiocyanate (Chirgwin *et al.*, 1979) and fractionated on an oligo(dT) column, was separated on a 1% formaldehyde-

agarose gel (Lehrach *et al.*, 1977) and transferred to nitrocellulose (Thomas, 1980). Rat kidney ribosomal RNA was used as size markers. Filters were hybridized and washed as described above.

#### Acknowledgements

We thank S.M. Blaine and S. Runciman for technical assistance; M. Shales for computer assistance; V. Auld and R. Dunn for the rat brain *igt11* library; H.F. Willard for the generous gift of genomic DNAs from human-mouse somatic cell hybrids; K. Letwin and T. Pawson for the gift of the rat *erk* cDNA probe; and C.J. Ingles, T. Pawson, J. Segall, and H.F. Willard for helpful discussions and critical review of this manuscript. This work was supported by the Medical Research Council of Canada.

#### References

- Allison, L.A., Moyle, M., Shales, M. & Ingles, C.J. (1985). *Cell*, **42**, 599-610.
- Auger, K.R., Serunian, L.A., Soltoff, S.P., Libby, P. & Cantley, L.C. (1989). *Cell*, **57**, 167-175.
- Auld, V.J., Goldin, A.L., Krafe, D.S., Marshall, J., Dunn, J.M., Catterall, W.A., Lester, H.A., Davidson, N. & Dunn, R.J. (1988). *Neuron*, **1**, 449-461.
- Breathnach, R. & Chambon, P. (1981). *Annu. Rev. Biochem.*, **50**, 349-383.
- Buchberg, A.M., Brownell, E., Nagata, S., Jenkins, N.A. & Copeland, N.G. (1989). *Genetics*, **122**, 153-161.
- Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. & Rutter, W.J. (1979). *Biochemistry*, **18**, 5294-5299.
- Edwards, A.M., Arquint, M., Braun, P.E., Roder, J.C., Dunn, R.J., Pawson, T. & Bell, J.C. (1988). *Mol. Cell. Biol.*, **8**, 2655-2658.
- Feinberg, A.P. & Vogelstein, B. (1983). *Anal. Biochem.*, **132**, 6-13.
- Hanks, S.K., Quinn, A.M. & Hunter, T. (1988). *Science*, **241**, 42-52.
- Hirai, H., Maru, Y., Hagiwara, K., Nishida, J. & Takaku, F. (1987). *Science*, **238**, 1717-1720.
- Hunter, T. & Cooper, J.A. (1985). *Ann. Rev. Biochem.*, **54**, 897-930.
- Kazlauskas, A., Ellis, C., Pawson, T. & Cooper, J.A. (1990). *Science*, **247**, 1578-1581.
- Konopka, J.B., Watanabe, S.M. & Witte, O.N. (1984). *Cell*, **37**, 1035-1042.
- Kornbluth, S., Paulson, K.E. & Hanafusa, H. (1988). *Mol. Cell. Biol.*, **8**, 5541-5544.
- Lawn, R.M., Fritsch, E.F., Parker, R.C., Blake, G. & Maniatis, T. (1978). *Cell*, **15**, 1157-1174.
- Lehrach, H., Diamond, D., Wozney, J.M. & Boedtker, H. (1977). *Biochemistry*, **16**, 4743-4751.
- Letwin, K., Yee, S.-P. & Pawson, T. (1988). *Oncogene*, **3**, 621-627.
- Lindberg, R.A. & Hunter, T. (1990). *Mol. Cell. Biol.*, **10**, 6316-6324.
- Lhotak, V., Greer, P., Letwin, K. & Pawson, T. (1991). *Mol. Cell. Biol.*, **11**, 2496-2502.
- Martin-Zanca, D., Hughes, S.H. & Barbacid, M. (1986). *Nature*, **319**, 743-748.
- Maru, Y., Hirai, H., Yoshida, M.C. & Takaku, F. (1988). *Mol. Cell. Biol.*, **8**, 3770-3776.
- Maru, Y., Hirai, H. & Takaku, F. (1990). *Oncogene*, **5**, 445-447.
- Meisenhelder, J., Suh, P.-G., Rhee, S.G. & Hunter, T. (1989). *Cell*, **57**, 1109-1122.
- Nairn, A.C., Hemmings, H.C. & Greengard, P. (1985). *Annu. Rev. Biochem.*, **54**, 931-976.
- Nudel, U., Zakut, R., Shani, M., Neuman, S., Levy, Z. & Yafe, D. (1983). *Nucleic Acids Res.*, **11**, 1759-1771.
- Roberts, W.M., Look, A.T., Roussel, M.F. & Sherr, C.J. (1988). *Cell*, **55**, 655-661.
- Sanger, F., Nicklen, S. & Coulson, A.R. (1977). *Proc. Natl. Acad. Sci. USA*, **74**, 5463-5467.
- Shier, P. & Watt, V.M. (1989). *J. Biol. Chem.*, **264**, 14605-14608.
- Shier, P., Willard, H.F. & Watt, V.M. (1990). *Cytogenet. Cell Genet.*, **54**, 80-81.
- Smart, J.E., Oppermann, H., Czernilofsky, A.P., Purchio, A.F., Erikson, R.L. & Bishop, J.M. (1981). *Proc. Natl. Acad. Sci. USA*, **78**, 6013-6017.
- Soares, M.B., Schon, E., Henderson, A., Kurathanasis, S.K., Cate, R., Zeitlin, S., Chirgwin, J. & Efstratiadis, A. (1985). *Mol. Cell. Biol.*, **5**, 2090-2103.
- Southern, E.M. (1975). *J. Mol. Biol.*, **98**, 503-517.
- Tabor, S. & Richardson, C.C. (1987). *Proc. Natl. Acad. Sci. USA*, **84**, 4767-4771.
- Thomas, P.S. (1980). *Proc. Natl. Acad. Sci. USA*, **77**, 5201-5205.
- Towbin, H., Staehelin, T. & Gordon, J. (1979). *Proc. Natl. Acad. Sci. USA*, **76**, 4350-4354.
- Ullrich, A. & Schlessinger, J. (1990). *Cell*, **61**, 203-212.
- Wahl, G.M., Stern, M. & Stark, G.R. (1979). *Proc. Natl. Acad. Sci. USA*, **76**, 3683-3687.
- Watt, V.M. & Willard, H.F. (1990). *Human Genet.*, **85**, 651-654.
- Yarden, Y. & Ullrich, A. (1988). *Annu. Rev. Biochem.*, **57**, 443-478.